Acceleration, Quantum, Information Technology and Algorithm Journal (AQILA)
Vol. 3, No. 1, June 2026, pp 35-40
ISSN: 3062-8555, DOI: 10.62123/aqila.v3i1.181

Deepfake Technology: Ethical Issues and Legal Gaps in Indonesian Cyber Law

Made Marshall Vira Deva[1*], Irfan Venny Rahmayanti[1], Intan Giri Anjani[1], Sutan Faiz Rasyid[1], Muharman Lubis[1]

[1]Telkom University, Bandung, Indonesia

*Corresponding Email: mademarshall97@gmail.com

DOI : 10.6213/aqila.v3i1.181

Received : June 15, 2026
Revised : June 20, 2026
Accepted : June 21, 2026

Keywords:
Cyber Ethics
Deepfake
Indonesian Cyber Law
Legal Gap
UU ITE

ABSTRACT

The rapid advancement of artificial intelligence (AI) has enabled the emergence of deepfake technology, which allows the manipulation of images, audio, and video to produce highly realistic yet fabricated content. In Indonesia, the proliferation of deepfakes poses significant ethical and legal challenges. This study examines the ethical implications of deepfake technology and identifies gaps in Indonesian cyber law, specifically within the Electronic Information and Transactions Law (UU ITE No. 11/2008 as amended by UU No. 19/2019 and UU No. 1/2024), the Pornography Law (UU No. 44/2008), and the Personal Data Protection Law (UU PDP No. 27/2022). Using a normative juridical research method with qualitative analysis of primary legal sources and secondary literature, this study finds that existing Indonesian legislation does not explicitly regulate deepfakes, creating a legal vacuum that leaves victims predominantly women without adequate legal protection. The findings highlight the urgent need for specific regulatory provisions addressing deepfake creation, distribution, and non-consensual intimate imagery (NCII). This paper concludes by proposing recommendations for legislative reform and ethical frameworks to guide both policymakers and technology users in Indonesia.

1. INTRODUCTION

Artificial intelligence (AI) has transformed digital communication in unprecedented ways. Among its most disruptive applications is deepfake technology a portmanteau of 'deep learning' and 'fake' which utilizes generative adversarial networks (GANs) to produce synthetic media that manipulates or replaces the likeness of real individuals [1]. Originally developed for entertainment and research purposes, deepfakes have increasingly been weaponized for malicious intent, including non-consensual pornography, political disinformation, identity fraud, and cyberbullying [2].

In Indonesia, the problem has grown substantially. A 2023 report by the Safe Internet Foundation (Yayasan Internet Aman) indicated that cases involving manipulated digital content including deepfake pornography have increased by over 70% compared to the previous year, with the majority of victims being women and public figures [3]. Despite this alarming trend, Indonesia lacks specific legislation targeting deepfake technology, creating a critical legal vacuum. This is corroborated by Komnas Perempuan's 2025 Annual Report (CATAHU), which recorded 376,529 cases of gender-based violence against women, a 14.07% increase from the previous year [28], indicating that online and AI-facilitated abuse forms part of a broader, worsening national pattern.

The existing legal framework, primarily UU ITE (Law No. 11/2008 as amended by Law No. 19/2019 and further revised by Law No. 1/2024), the Pornography Law (Law No. 44/2008), and the newly enacted Personal Data Protection Law (UU PDP No. 27/2022), provides only indirect and fragmented protections. These laws were enacted before the widespread adoption of generative AI, rendering them insufficient to address the technical specificity of deepfake-based offenses [4].

This study aims to: (1) analyze the ethical dimensions of deepfake technology in the context of cyber ethics; (2) identify specific legal gaps in Indonesian cyber law regarding deepfakes; and (3) propose normative recommendations for legislative reform. By bridging legal analysis with ethical frameworks, this research contributes to the growing body of scholarship on AI governance and digital rights in Indonesia

2. LITERATURE REVIEW

2.1. Deepfake Technology and Generative AI

Deepfakes are synthetic media generated using deep learning techniques, particularly GANs introduced by Goodfellow et al. [5]. A GAN consists of two neural networks a generator and a discriminator that compete to produce increasingly realistic outputs. The generator creates fake content, while the discriminator attempts to distinguish it from real content; this adversarial process results in highly convincing synthetic media [6].

Deepfake generation encompasses several distinct technical approaches with differing legal-evidentiary implications. Face-swapping replaces one individual's facial identity onto another's body in existing footage, typically used in NCII cases. Voice cloning synthesizes an individual's speech patterns from limited audio samples, increasingly used in fraud and impersonation scams[7]. Full-body synthesis generates an entirely fabricated person or recreates a subject's full body movement, raising distinct questions about whose biometric data was actually processed. These distinctions matter legally because the type of synthesis determines what digital evidence (facial landmarks, voice spectrograms, motion data) must be examined during prosecution [8].

Each of these synthesis methods technically requires processing an individual's biometric data face geometry for face-swapping, vocal biometric signatures for voice cloning which falls squarely within the definition of "specific personal data" under UU PDP Article 4, requiring explicit consent. This creates a direct, but currently unenforced, link between the technical process of deepfake creation and Indonesia's data protection regime. A further technical challenge lies in detection and forensic verification. Current deepfake detection methods, such as artifact analysis and digital watermarking, face an inherent limitation: detection techniques are developed reactively, while generation techniques improve continuously, creating a persistent "arms race" dynamic [9]. Watermarking, while promising, can be stripped or degraded through re-compression, and no Indonesian forensic standard currently exists for authenticating synthetic media as courtroom evidence [10]. This technical gap directly explains the "no technical standard for synthetic media evidence" finding in Table 1, and raises practical questions about how Indonesian courts would evaluate deepfake-based evidence absent a recognized verification protocol.

Early applications of deepfakes were largely benign, including face-swapping in films and voice cloning for accessibility tools [8]. However, the democratization of AI tools including readily available applications such as FaceSwap, DeepFaceLab, and web-based deepfake generators has lowered the technical barrier to misuse [11]. Research by Chesney and Citron [2] classified deepfakes as a 'liar's dividend' threat: even the mere possibility of fabricated content undermines public trust in authentic media.

2.2. Ethical Dimensions of Deepfakes

The ethics of deepfake technology can be analyzed through multiple frameworks. From a consequentialist perspective, deepfakes cause measurable harm: psychological trauma to victims, erosion of democratic discourse, and economic damage through fraud [12]. From a deontological standpoint, the non-consensual use of an individual's likeness violates their autonomy and dignity, regardless of the consequences [13].

Floridi et al. [14] propose an AI ethics framework centered on five principles: beneficence, non-maleficence, autonomy, justice, and explicability. Applied to deepfakes, the technology consistently fails the non-maleficence and autonomy tests when deployed without consent. The concept of 'digital dignity' an individual's right to control their digital identity is central to cyber ethics discourse and is directly threatened by deepfakes [15]. In the Indonesian context, Rahardjo [16] argues that the cultural emphasis on collective harmony (gotong royong) does not diminish individual rights in cyberspace; rather, it demands greater communal responsibility in combating technology-facilitated abuse. This perspective is critical for formulating culturally appropriate responses to deepfake misuse.

2.3. Legal Frameworks for Deepfakes Globally

Several jurisdictions have begun addressing deepfakes legislatively. In the United States, the DEEPFAKES Accountability Act (proposed 2019, 2021) sought to require disclosure watermarks on synthetic media [17]. China issued Regulations on Deep Synthesis Internet Information Services effective January 2023, requiring labeling of AI-generated content and prohibiting non-consensual deepfakes [18]. The United Kingdom's Online Safety Act 2023 criminalizes the sharing of intimate deepfakes without consent [19].

The European Union, under the AI Act (2024), classifies systems used for real-time biometric identification and certain synthetic media generation as high-risk or prohibited AI applications, mandating transparency requirements [20]. These international developments illustrate a growing consensus that deepfakes require targeted regulation beyond general cybercrime statutes.

2.4. Indonesian Legal Framework for Cyber Offenses

Indonesia's primary cyber law, UU ITE (Law No. 11/2008), has undergone two revisions (2019 and 2024). Key provisions potentially applicable to deepfake offenses include Article 27(1) on content violating decency norms, Article 27(3) on defamation, Article 28(1) on false information causing consumer harm, and Article 35 on data falsification [4]. However, these provisions were not designed with AI-generated synthetic media in mind and fail to capture the technical nuances of deepfakes.

The Pornography Law (UU No. 44/2008) prohibits production, distribution, and possession of pornographic material, including that involving manipulation. Article 4(1) could potentially be applied to deepfake pornography, yet the law does not address the concept of non-consensual intimate imagery (NCII) or synthetic identity exploitation [21].

UU PDP (Law No. 27/2022) introduces data subject rights including image rights and biometric data protections. Article 4 classifies biometric data as specific personal data requiring explicit consent for processing. The use of an individual's facial data to generate deepfakes without consent may constitute a violation of UU PDP; however, enforcement mechanisms remain unclear as the regulatory body (Lembaga PDP) has not yet been fully established [22].

3. RESEARCH METHODS

This study employs a normative juridical (yuridis normatif) research method, a standard approach in Indonesian legal scholarship for analyzing the consistency, coherence, and completeness of existing legal norms [23]. The normative approach is appropriate here because the primary objective is to identify gaps in positive law rather than to measure empirical phenomena.

The research framework consists of three stages, as illustrated in Figure 1. In Stage 1, primary legal sources are collected and analyzed: these include the full texts of UU ITE (Law No. 11/2008 jo. Law No. 19/2019 jo. Law No. 1/2024), UU Pornografi (Law No. 44/2008), UU PDP (Law No. 27/2022), and relevant Constitutional Court decisions. In Stage 2, secondary sources are reviewed, comprising peer-reviewed journal articles, legal commentaries, government reports, and comparative international legislation. In Stage 3, a gap analysis is performed by systematically mapping deepfake-related harms against existing legal provisions to identify coverage and lacunae.

The qualitative analysis applies three interpretive lenses: (1) grammatical interpretation examining the literal text of statutory provisions; (2) systematic interpretation analyzing provisions in relation to the broader legal system; and (3) teleological interpretation assessing whether provisions fulfill their intended policy objectives in the context of AI-generated synthetic media [24].

Data validity is ensured through source triangulation: legal texts are cross-referenced with legislative histories (risalah sidang), academic commentaries, and comparative international legal analysis. Ethical clearance is not required for normative legal research as no human subjects are involved.

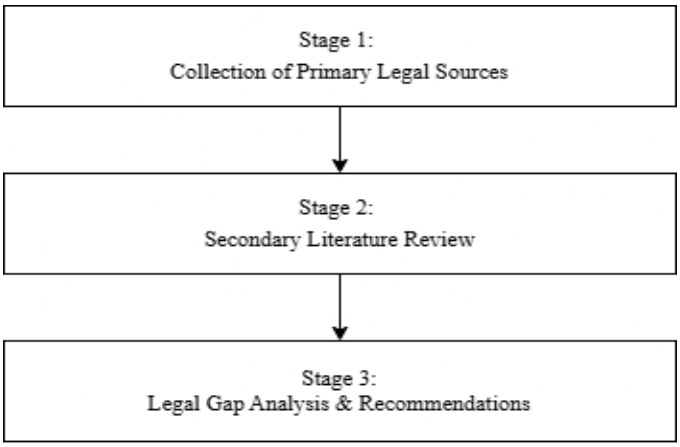

Figure 1. Research Framework

4. DISCUSSION AND RESULT

4.1. Ethical Analysis of Deepfakes in Indonesia

The ethical analysis reveals that deepfake technology in Indonesia operates in a context where both technological literacy and legal awareness among users remain limited. Three primary ethical violations are identifiable:

First, violation of personal autonomy and digital dignity. When an individual's biometric data specifically facial features is used without consent to generate synthetic media, it constitutes a fundamental breach of their right to control their own identity [15]. This is particularly acute in cases of deepfake pornography, where victims report severe psychological harm including depression, anxiety, and social withdrawal [25]. This ethical violation maps directly onto the legal vacuum identified in the "Deepfake NCII / pornography" row of Table 1, where no provision protects victim identity once synthetic content is distributed.

Second, epistemic harm and erosion of public trust. Deepfakes deployed for political disinformation corrupt the information environment that democratic deliberation depends upon. In Indonesia's politically charged media landscape, AI-generated videos mimicking public figures pose a direct threat to electoral integrity [26].

Third, gendered harm patterns. Research consistently shows that approximately 96% of deepfake content online is non-consensual pornography, and the overwhelming majority of victims are women [27]. This gendered dimension necessitates that

any ethical and legal response explicitly address the intersection of technology, gender, and power. This gendered pattern directly informs Recommendation 2 in Section 4.4, which calls for amending the Pornography Law to center victim consent rather than content authenticity.

4.2. Legal Gap Analysis

Table 1 presents a systematic mapping of deepfake harm categories against applicable Indonesian legal provisions, revealing significant gaps. This gap is further compounded by Constitutional Court Decision No. 105/PUU-XXII/2024 (29 April 2025) [29], which narrowed Article 27A of UU ITE to apply only to natural persons, excluding institutions and corporations from filing defamation claims. This means that synthetic media defaming an organization or institution as opposed to an individual currently falls entirely outside the scope of Indonesia's primary defamation provision.

Table 1. Mapping of Deepfake Harms to Indonesian Legal Provisions

Harm Category	Applicable Provision	Gap Identified
Deepfake NCII / pornography	UU Pornografi Art. 4(1); UU ITE Art. 27(1)	No explicit provision for AI-synthesized NCII; victim identity protection absent
Identity falsification / fraud	UU ITE Art. 35	Does not address AI-generated biometric manipulation
Defamation via synthetic video	UU ITE Art. 27(3)	Broad application; no technical standard for synthetic media evidence
Biometric data misuse	UU PDP Art. 4; Art. 65	Enforcement body (Lembaga PDP) not yet operational; liability unclear
Political deepfake disinformation	UU ITE Art. 28(1)	Requires proof of financial harm; political harm threshold unaddressed
Deepfake involving minors	UU PA; UU Pornografi Art. 4(2)	Partial coverage; synthetic child imagery definition unclear

The gap analysis in Table 1 demonstrates that while Indonesian law offers partial remedies through analogical interpretation, none of the existing provisions explicitly target deepfake technology. This creates legal uncertainty for both victims seeking redress and law enforcement agencies attempting prosecution.

A critical structural gap is the absence of a consent-based framework for synthetic identity use. Unlike UU PDP which requires explicit consent for biometric data processing, UU ITE and the Pornography Law do not establish a similar consent standard for the use of an individual's digital likeness in AI-generated media. This legislative gap is particularly problematic because deepfake generation technically involves the processing of biometric data bringing it within UU PDP's scope yet the punitive mechanisms of UU PDP are weaker than those of UU ITE.

4.3. Comparative Analysis

Table 2 provides a comparative overview of deepfake regulation across selected jurisdictions, contextualizing Indonesia's legal position. The comparative analysis reveals that Indonesia lags significantly behind leading jurisdictions in addressing deepfakes legislatively. While China and the EU have adopted proactive, technology-specific regulatory frameworks, Indonesia relies on analogical application of pre-AI legislation. This approach is inadequate given the technical complexity of deepfake detection and prosecution.

Table 2. Comparative Deepfake Regulation

Jurisdiction	Key Regulation	Approach
China	Deep Synthesis Regulations 2023	Mandatory labeling; consent requirement; platform liability
United Kingdom	Online Safety Act 2023	Criminalization of NCII deepfake sharing without consent
European Union	AI Act 2024	Risk-based classification; transparency obligations
United States	State-level laws (CA, TX, VA)	NCII criminalization; election deepfake disclosure
Indonesia	UU ITE, UU Pornografi, UU PDP	No specific deepfake provision; analogical application only

4.4. Proposed Recommendations

The recommendations below are derived directly from the ethical violations identified in Section 4.1 and the legal gaps mapped in Table 1. Specifically: the violation of autonomy and digital dignity (4.1) corresponds to the absence of victim-identity

protection in the Pornography Law (Table 1, row 1), addressed through Recommendation 1 and 2; the epistemic harm from political disinformation (4.1) corresponds to the unaddressed political-harm threshold in UU ITE Article 28(1) (Table 1, row 5), addressed through Recommendation 4; and the gendered harm pattern (4.1) corresponds to the consent-standard gap (4.2), addressed through Recommendation 2.

Based on the gap analysis and comparative findings, this study proposes the following normative recommendations. First, enactment of specific deepfake provisions within the UU ITE framework. A dedicated article should define 'synthetic digital media' and 'AI-generated identity manipulation,' establishing criminal liability for non-consensual creation and distribution of deepfakes. The definition should be technology-neutral to accommodate future AI developments. Second, amendment of the Pornography Law to explicitly include non-consensual intimate imagery (NCII) regardless of whether the content is authentic or AI-generated. The concept of victim consent should be central to determining criminality, aligning Indonesian law with international best practices [19]. Third, acceleration of UU PDP implementation. The immediate establishment of the Lembaga PDP (Personal Data Protection Authority) is essential. This body should issue specific technical guidelines on biometric data use in AI applications, including explicit prohibition of facial data processing for deepfake generation without consent.

Fourth, adoption of mandatory disclosure requirements. Platforms operating in Indonesia should be required to label AI-generated content, drawing from China's deep synthesis labeling model and the EU AI Act's transparency obligations. Fifth, development of a National Cyber Ethics Framework. Beyond legal sanctions, Indonesia requires a comprehensive cyber ethics education program integrated into school curricula and university programs that addresses AI ethics, digital consent, and responsible technology use.

5. CONCLUSION

This study has demonstrated that deepfake technology presents significant ethical challenges and exploits critical gaps in Indonesian cyber law. The analysis of UU ITE, the Pornography Law, and UU PDP reveals that while partial legal remedies exist through analogical interpretation, Indonesia lacks specific legislation targeting AI-generated synthetic media. This legal vacuum leaves victims disproportionately women without adequate protection and creates enforcement uncertainty.

The comparative analysis confirms that Indonesia lags behind China, the United Kingdom, and the European Union in deepfake governance. The absence of consent-based frameworks, mandatory disclosure requirements, and a functional personal data protection authority compounds the legislative inadequacy.

This research recommends: (1) enacting dedicated deepfake provisions within UU ITE; (2) amending the Pornography Law to cover NCII regardless of authenticity; (3) expediting UU PDP enforcement infrastructure; (4) adopting AI-content labeling mandates; and (5) developing a national cyber ethics education framework. Future research should examine empirical case studies of deepfake-related prosecutions in Indonesia and evaluate the effectiveness of proposed legislative interventions upon enactment

REFRENCES

[1] I. Goodfellow *et al.*, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.

[2] B. Chesney and D. Citron, "Deep fakes: A looming challenge for privacy, democracy, and national security," *Calif. Law Rev.*, vol. 107, no. 6, pp. 1753–1820, 2019, doi: 10.15779/Z38RV0D15J.

[3] D. Nainggolang, K. I. Tangkulung, L. M. B. Siregar, and Z. L. Lampa, "Strategi Pengawasan Konten Digital dalam Melindungi Anak dari Paparan Konten Berbahaya di Media Sosial," *Jurnal Pengabdian Masyarakat dan Riset Pendidikan*, vol. 4, no. 3, pp. 14575–14579, Dec. 2025, doi: 10.31004/jerkin.v4i3.4427.

[4] G. E. Sidi Artama and N. P. Ega Parwati, "ANALISIS YURIDIS TERHADAP PENYALAHGUNAAN KECERDASAN BUATAN DALAM PENIPUAN BERMODUS PENCULIKAN ANAK MELALUI IMITASI SUARA," 2023.

[5] I. J. Goodfellow *et al.*, "Generative Adversarial Nets," 2019. [Online]. Available: http://www.github.com/goodfeli/adversarial

[6] T. Karras, S. Laine, and Aila Timo, "A Style-Based Generator Architecture for Generative Adversarial Networks Timo Aila NVIDIA," 2018. [Online]. Available: https://github.com/NVlabs/stylegan

[7] B. Zhang, H. Cui, V. Nguyen, and M. Whitty, "Audio Deepfake Detection: What Has Been Achieved and What Lies Ahead," *Sensors*, vol. 25, no. 7, Apr. 2025, doi: 10.3390/s25071989.

[8] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "Deepfakes and beyond: A Survey of face manipulation and fake detection," *Information Fusion*, vol. 64, pp. 131–148, Dec. 2020, doi: 10.1016/j.inffus.2020.06.014.

[9] Y. Mirsky and W. Lee, "The Creation and Detection of Deepfakes," Jul. 31, 2021, *Association for Computing Machinery*. doi: 10.1145/3425780.

[10] T. Wang, X. Liao, K. P. Chow, X. Lin, and Y. Wang, "Deepfake Detection: A Comprehensive Survey from the Reliability Perspective," Oct. 2024, doi: 10.1145/3699710.

[11] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "FaceForensics++: Learning to Detect Manipulated Facial Images," 2019.

[12] H. Farid, "Creating, Using, Misusing, and Detecting Deep Fakes," *Journal of Online Trust and Safety*, vol. 1, no. 4, Sep. 2022, doi: 10.54501/jots.v1i4.56.

[13] L. Lazard, R. Capdevila, E. L. Turley, K. Gilfoyle, and N. Stavropoulou, "Deepfake Technology and Gender-Based Violence: A Scoping Review," 2025, *SAGE Publications Ltd.* doi: 10.1177/15248380251384271.
[14] L. Floridi *et al.*, "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations," *Minds Mach. (Dordr).*, vol. 28, no. 4, pp. 689–707, Dec. 2018, doi: 10.1007/s11023-018-9482-5.
[15] Supriandi, Khairunnisa, and W. Utama Putra, "Hak Asasi Manusia di Ranah Digital: Analisis Hukum Siber dan Kebebasan Online," 2024.
[16] M. Hera Yulianto, C. Hasanudin, and E. Duwi Saputri, "Upaya Meningkatkan Gotong Royong dan Kerukunan Mahasiswa di Era Society 5.0," 2025.
[17] A. Shanaz Joan Parsan *et al.*, "Unmasking Deepfakes A Review of Technology, Regulation, Challenges and Policy Implications." [Online]. Available: www.szc-group.com
[18] M. Zou and L. Zhang, "Navigating China's regulatory approach to generative artificial intelligence and large language models," *Cambridge Forum on AI: Law and Governance*, vol. 1, 2025, doi: 10.1017/cfl.2024.4.
[19] J. Ward, "The Weaponisation of Artificial Intelligence (AI): Legal Shortfalls and Regulatory Difficulties in Governing Non-Consensual Intimate Deepfakes (NCIDs)," Jul. 21, 2025. doi: 10.20944/preprints202507.1645.v1.
[20] S. Arda, "Taxonomy to Regulation: A (Geo)Political Taxonomy for AI Risks and Regulatory Measures in the EU AI Act," Apr. 2024, [Online]. Available: http://arxiv.org/abs/2404.11476
[21] F. Ivana Putri, N. Soekorini, and M. Taufik, "Aspek Tindak Pidana Penyebaran Konten Asusila LGBT (Lesbian, Gay, Biseksual, Transgender) di Media Sosial," *Jurnal Sosial dan Teknologi (SOSTECH)*, vol. 6, 2026.
[22] S. Nurkholisah, D. Rismana, A. E. Nugroho, A. Munjiyah, and Q. Ayunisa, "Tantangan Kriminalisasi Deepfake dalam Hukum Pidana Indonesia Challenges in Criminalizing Deepfakes under Indonesian Criminal Law," *Jurnal USM Law Review*, vol. 8, 2025, doi: 10.2139/ssrn.4321456.
[23] P. M. Marzuki, "Pengantar Ilmu Hukum," 2008.
[24] S. Sigit *et al.*, "METODOLOGI RISET HUKUM," 2020.
[25] L. Carmona, D. Dasgupta, C. Noell Birondo, E. Day, and S. L. Crites, "AI-FACILITATED ABUSE & MISOGYNY: A FEMINIST RESPONSE TO NON-CONSENSUAL INTIMATE IMAGE DEEPFAKES," 2025.
[26] C. Wardle and H. Derakhshan, "INFORMATION DISORDER : Toward an interdisciplinary framework for research and policy making Information Disorder Toward an interdisciplinary framework for research and policymaking," 2017. [Online]. Available: www.coe.int
[27] D. S. Louk, "Deepfakes, Real Enforcement Challenges," 2026. [Online]. Available: https://web.archive.org/web/20250124021021/https://inhope.org/EN/articles/what-is-ncii.
[28] Komnas Perempuan, "Siaran Pers: Peluncuran Catatan Tahunan Kekerasan terhadap Perempuan 2025," 2026. [Online]. Available: https://komnasperempuan.go.id/siaran-pers-detail/siaran-pers-komnas-perempuan-peluncuran-catatan-tahunan-kekerasan-terhadap-perempuan-2025.
[29] Mahkamah Konstitusi Republik Indonesia, "MK Mempertegas Pemaknaan Unsur-Unsur Pencemaran Nama Baik dalam UU ITE," Putusan No. 105/PUU-XXII/2024, 2025. [Online]. Available: https://www.mkri.id/berita/-23133.